Faculty of Cognitive Science and Human Development

DEVELOPMENT AND IMPLEMENTATION OF FUZZY ARTMAP NEURAL NETWORK MODEL IN THE PATTERN CLASSIFICATION OF PRODUCT DESIGN PREFERENCES: A CASE STUDY FOR EYEGLASSES FRAME DESIGN

Tang Weng Chin

UNIVERSITI MALAYSIA SARAWAK
2003

UNIVERSITI MALAYSIA SARAWAK

BORANG PENGESAHAN STATUS TESIS

JUDUL: DEVELOPMENT AND IMPLEMENTATION OF FUZZY ARTMAP NEURAL NETWORK MODEL IN THE PATTERN CLASSIFICATION OF PRODUCT DESIGN PREFERENCES: A CASE STUDY FOR EYEGLASSES FRAME DESIGN

SESI PENGAJIAN: 2002/2003

Saya TANG WENG CHIN

mengaku membenarkan tesis * ini disimpan di Pusat Khidmat Maklumat Akademik, Universiti Malaysia Sarawak dengan syarat-syarat kegunaan seperti berikut:

1. Tesis adalah hakmilik Universiti Malaysia Sarawak
2. Pusat Khidmat Maklumat Akademik, Universiti Malaysia Sarawak dibenarkan membuat salinan untuk tujuan pengajian sahaja
3. Pusat Khidmat Maklumat Akademik, Universiti Malaysia Sarawak dibenarkan membuat pendigitan untuk membangunkan Pangkalan Data Kandungan Tempatan
4. Pusat Khidmat Maklumat Akademik, Universiti Malaysia Sarawak dibenarkan membuat salinan tesis ini sebagai bahan pertukaran antara institusi pengajian tinggi
5. ** sila tandakan (✓)

☐ SULIT (mengandungi maklumat yang berdarjah keselamatan atau kepentingan seperti yang termaktub di dalam AKTA RAHSIA RASMI 1972)

☐ SULIT (mengandungi maklumat Terhad yang ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)

☑ TIDAK TERHAD

Disahkan oleh

(TANDATANGAN PENULIS)

(TANDATANGAN PENYELIA)

Alamat Tetap:
160, Juru Village,
14000 Bukit Mertajam,
Pulau Pinang, Malaysia.

Tarikh:

Tarikh: 29/03/03.

Catatan: * Tesis yang dimakudkan sebagai tesis bagi Ijazah Doktor Falsafah, Sarjana dan Sarjana Muda
Jika tesis ini SULIT atau TERHAD, sila lampirkan surat daripada pihak berkuasa/organisasi berkenaan
Dengan menyatakan sekali sebab dan tempoh tesis ini perlu dikelaskan sebagai SULIT atau TERHAD

DEVELOPMENT AND IMPLEMENTATION OF FUZZY ARTMAP NEURAL NETWORK MODEL IN THE PATTERN CLASSIFICATION OF PRODUCT DESIGN PREFERENCES: A CASE STUDY FOR EYEGLASSES FRAME DESIGN

by

Tang Weng Chin

This project is submitted in partial fulfilment of the requirement for a
Bachelor of Science (Honours) in Cognitive Science
Faculty of Cognitive Sciences and Human Development,
Universiti Malaysia Sarawak

The project entitled Development and Implementation of Fuzzy ARTMAP Neural Network Model in the Pattern Classification of Product Design Preferences: A Case Study for Eyeglasses Frame Design was prepared by Tang Weng Chin and submitted to the Faculty of Cognitive Sciences and Human Development in partial fulfilment of the requirement for a Bachelor of Science (Honours) in Cognitive Science.

Received for examination by:

(Mr. Philip Nuli Anding)

Date:

29/03/03

ACKNOWLEDGEMENT

I would like to express my appreciation to Mr. Philip Nuli Anding for his help and guidance during my efforts of completing this thesis.

I also like to express my gratitude to Mr. Teh Chee Siong for his valuable idea, suggestion and guidance during my effort of developing this project.

I would also like to thank all my lecturers, who have imparted valuable knowledge to me through their lectures and discussions, thus equipping me with the necessary skills to complete this project.

Other than that, special thanks to Ms. Chan Sok Feng for her support and help throughout the process of completing this thesis.

Last, but not least, I would like to express my thanks to my family and friends for their understanding and support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

DEVELOPMENT AND IMPLEMENTATION OF FUZZY ARTMAP NEURAL NETWORK MODEL IN THE PATTERN CLASSIFICATION OF PRODUCT DESIGN PREFERENCES: A CASE STUDY FOR EYGLASSES FRAME DESIGN

Tang Weng Chin

This project is concerned with the development and implementation of Fuzzy ARTMAP Neural Network model in the pattern classification of product design preferences. In product design, it is a difficult task to classify the user's preferences according the physical design elements. Conventional methods such as survey are normally used to perform the pattern classification task. However, these methods normally limit the identification of preferences pattern in the existing products and require statistical knowledge. Thus it would be beneficial if a computerized pattern classification system could be utilized to assist the product designer to identify the user's preferences towards product design. The artificial neural network (ANNs) model of Fuzzy Adaptive Resonance Theory Mapping (Fuzzy ARTMAP) was used in this project to perform the pattern classification task in the product design domain. This model was coded in C/C++ language based on the original architecture and learning algorithm proposed by Carpenter and Grossberg. The pattern classification performance of this model was evaluated using two sets of universal accepted datasets that is the Pima Indian Diabetes dataset and the Wine Classification dataset. A data collection tool was developed in this project using the combination of Java programming language, Virtual Reality Modelling Language (VRML) and External Authoring Interface (EAI). This data collection tool provides the facility of generating virtual product model based on the adjustment of the design parameters. Eyeglasses frame design was taken as a case study for this project. The data of the user's preferences towards eyeglasses frame design were captured using the data collection tool. The Fuzzy ARTMAP model was trained using these data and this model was integrated into the implementation system. Finally, the implementation system will be able to classify and predict the target user based on the configuration of eyeglasses frame design.

ABSTRAK

PEMBANGUNAN DAN IMPLEMENTASI MODEL RANGKAIAN NEURAL BUATAN FUZZY ARTMAP DALAM MENGKLASIFIKASIKAN CORAK KEGEMARAN UNTUK REKABENTUK PRODUK: SATU KES KAJIAN UNTUK REKABENTUK RANGKA CERMIN MATA

Tang Weng Chin

Projek ini adalah mengenai pembangunan dan implementasi sebuah model Rangkaian Neural Buatan yang dikenali sebagai Fuzzy ARTMAP dalam tugas mengklasifikasikan corak kegemaran pengguna terhadap rekabentuk produk. Tugas mengklasifikasikan corak kegemaran pengguna terhadap rekabentuk sesuatu produk adalah susah dilakukan dalam bidang rekabentuk produk. Teknik-teknik yang lazimnya digunakan untuk melakukan tugas ini adalah seperti teknik soal-selidik. Teknik-teknik ini mempunyai kelemahan dari segi menghadkan menemuan corak kegemaran terhadap produk yang sedia-ada dalam pasaran serta memerlukan pengetahuan statistik untuk menggunakan teknik-teknik ini. Dengan itu, sebuah sistem komputer diperlukan untuk membantu para pereka dalam melakukan tugas mengklasifikasikan corak kegemaran pengguna terhadap rekabentuk sesuatu produk. Model rangkaian neural buatan yang digunakan dalam projek ini dikenali sebagai Fuzzy Adaptive Resonance Theory Mapping (Fuzzy ARTMAP). Model in dikodkan dengan menggunakan bahasa pengaturcara C/C++ berdasarkan kepada arkitektur dan algoritma pembelajaran yang dicadangkan oleh Carpenter dan Grossberg. Keupayaan model in dalam tugas mengklasifikasikan corak diuji oleh dua set data iaitu data-data Pima Indian Diabetes dan data-data Wine Classification. Satu sistem pengumpulan data turut dibangunkan dengan menggunakan kombinasi bahasa pengaturcaraan Java, Virtual Reality Modelling Language (VRML) dan External Authoring Interface (EAI). Sistem pengumpulan data ini membolehkan suatu model produk maya dihasilkan melalui pengubahan pada parameter-parameter rekabentuk produk. Rekabentuk rangka cermin mata telah dipilih sebagai kajian kes dalam projek ini. Data-data mengenai kegemaran pengguna terhadap rekabentuk rangka cermin mata dikumpul dengan menggunakan sistem pengumpulan data tersebut. Model Fuzzy ARTMAP akan dilatih dengan menggunakan data-data tersebut dan model yang terlatih ini akan diintegrasikan ke dalam sistem implementasi. Sistem implementasi ini berupaya mengklasifikasikan serta berupaya meramal kumpulan pengguna berdasarkan kepada rekabentuk rangka cermin mata.

CHAPTER ONE
INTRODUCTION

1.1 Background

Artificial neural network (ANN) is one of the branches of artificial intelligence (AI). It is also known as connectionist system, parallel distributed processing (PDP), neural computing and artificial neural system.

Artificial neural network tries to mimic how the brain and nervous system works. It was inspired by the studies of various disciplines includes neurosciences and psychology. Artificial neural network is famous for its inductive learning ability that is the ability of learning from examples. Besides that, it is also able to generalize to previously unseen data, to extract information without specifying a data model and the *train-and-go* characteristic.

Artificial neural network is well known for its ability in tackling problems in pattern classification, pattern recognition, fault detection, medical diagnosis, business and others (Tan, 2001). However, artificial neural network is still a new approach in the product design domains. Recently artificial neural network technique was used in the product design field to capture the pattern of customer preferences, the pattern of customer feelings towards a product design and others. Such a technique provides an alternative way to capture the preferences of customers for the designer to cope with the current product design trend, which is "listen to voice of customer" (Hsiao & Huang, 2001).

1.2 Problem Statement

Pattern classification is a process of assigning an input pattern into one of the targeted classes according to some decision rules. This is a daily cognition process carried out by the human and plays a crucial role in decision-making. For example, doctor classifying diseases based on the patient's symptoms, weather forecaster predict weather based on the cloud's attributes, mechanic determines fault in machine based on it's condition. People constantly receive information from their environment and have to make decision based on the patterns of the information received. When encountered with pattern classification problem, then the stored knowledge and past experience can be used to assist in decision-making.

However, in product design, it is a difficult task to classify the user preferences into the correct classes. Survey is the common method used to identify user preferences toward certain product's physical features. But, this conventional method has its constraint, which limits the identification of user preferences only at the existing product on the market. Besides that, statistical knowledge is required to perform data analysis for the identification of user preferences.

Despite that, a computerized pattern classification system is needed to assist product designer to identify user preferences classes. This computerized system must be able to classify user preferences into the targeted classes and also to make prediction on user preferences based on the physical features of the product design. To do so, this system must be able to make use of the stored information. Therefore, the abilities of learning, a modification process of stored information through experience would be an essential characteristic of this system.

Hence, ANN methods can be applied in this computerized system due to its strength in machine learning, inductive ability and its well-proven abilities in pattern recognition and pattern classification.

1.3 Objective and Scope

The objective of this project is to develop a computerized pattern classification system using ANN. The main focus in this study is to investigate the role of ANN as a technique of pattern classification in product design domain. There are four specific objectives included to this project.

a) To study both the theoretical and algorithmic aspect of an ANN model that is the Fuzzy Adaptive Reasoning Theory Mapping (Fuzzy ARTMAP).
b) To develop the Fuzzy ARTMAP model based on the original Fuzzy ARTMAP model proposed by Stephen Grossberg and Gail Carpenter.
c) To evaluate the performance of the Fuzzy ARTMAP model by using two sets of universal accepted datasets.
d) To implement the Fuzzy ARTMAP model into the pattern classification of product design preferences.

The scope of this study focussed on a specific ANN model mentioned above, the Fuzzy ARTMAP model.

1.4 Significance of Research

This research is significant for the investigation of the potential of Fuzzy ARTMAP model in pattern classification of product design preferences. Besides that, this research is also important for providing an alternative solution for the pattern classification problem in the product design domain.

1.5 Summary

There are all together five chapters in this project reports. Chapter one is the introduction to this project. In this chapter, the problem statement and the objective of this project are discussed. Chapter two presents some literature reviews related to this study. For instance, ANN is introduced and this included the definition of ANN and the reviews of both the biological and historical aspects of ANN. This is then followed by the discussion of various ANN models in the Adaptive Resonance Theory (ART) models. Both the architecture and learning algorithm of each ART models are discussed. Chapter three covers the methodology used for this project such as problem identification, Fuzzy ARTMAP model development and implementation. Chapter four discuss about the results of this project. Conclusion and suggestion of further work will be in chapter five.

CHAPTER TWO
LITERITURE REVIEW

2.1 Definition of Artificial Neural Network

Artificial Neural Network is an information-processing paradigm that is inspired by the studies of the brain and nervous system. It is composed of a large number of highly interconnected processing elements or neurons working in unison to solve specific problems. It is configured for a specific application, such as pattern recognition or data classification, through a learning process.

There is no generally accepted definition of what constitutes an artificial neural network. This is some of the definition of artificial neural network.

"... a neural network is a system composed of many simple processing elements operating in parallel whose function is determined by network structure, connection strengths, and the processing performed at computing elements or nodes." (Darpa, 1988)

"A neural network is a massively parallel distributed processor that has a natural propensity for storing experiential knowledge and making it available for use. It resembles the brain in two respects:
Knowledge is acquired by the network through a learning process. Interneuron connection strengths known as synaptic weights are used to store the knowledge." (Haykin, 1994)

"An artificial neural network is an information processing system that has certain performance characteristics in common with biological neural network." (Faussett, 1994)

2.2 Historical Background of Artificial Neural Network

Artificial neural network appear to be a recent development. However, the field was established before the advent of computers, and has survived at least one major setback and several eras. The history of artificial neural network can be divided into several periods.

The first attempt at artificial neural network was in the 1940s, when American scientists Warren McCulloch and Walter Pitts develops models of artificial neural network based on their understanding in neurology. Their networks were based on simple neurons, which were considered to be binary devices with fixed thresholds. This model was made up of binary decision units (BDNs) and showed that it could perform any logical function on its inputs. The results showed that the BDN was a simple model that was similar to the nerve cell used in the human brain to support thinking.

The attempt in artificial neural network was followed by the Perceptron model developed by another American scientist, Frank Rosenblatt and several of his colleagues in 1958. The Perceptron had three layers with the middle layer known as the association layer. This system could learn to connect or associate a given input to a random output unit. This model was able to solve some simple classification problems. Another system was the ADALINE (*ADAptive*

Linear Element) which was developed in 1960 by Widrow and Hoff of Stanford University. The ADALINE was an analogue electronic device made from simple components. The method used for learning was different to that of the Perceptron, it employed the Least-Mean-Squares (LMS) learning rule.

The dark age of artificial neural network began in 1969, when Marvin Minsky and Seymour Papert proof the limitation of the perceptron model. They showed that the perceptron was not able to solve some very simple pattern classification task, such as separating binary patterns (0,0), (1,1) from the patterns (1,0), (0,1) which was known as the parity problem, or the XOR problem. To solve such a problem, it was necessary to have neurons whose output were not available to the outside world and such neurons are known as hidden neurons. As a result, the attention and attraction towards the field of artificial neural network was set-back.

The reawakening era of artificial neural network began when the difficulty of the training hidden neurons was solved by the back-propagation algorithm, originally introduced by Paul Werbos (1974) and independently discovered by Parker (1985) and LeCun (1985). The back-propagation was then highly publicized by the PDP Group of Rumelhart and McClelland (1986). The reawakening era followed by other network models includes the Hopfield network, Kohonen network and Carpenter and Grossberg Adaptive Resonance Network.

2.3 Biological Neural Network

According to Arbib (1997), the human brain consists of many different types of neurons. Each neuron has three main components, which are the dendrites, the soma (cell body) and a long fiber called the axon. The synapses are chemical junctions between axons and dendrites and between axons of different neurons in which the transmission of information in the form of neurotransmitters or chemical ions occur.

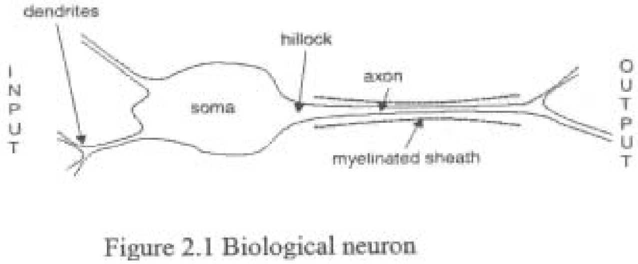


Figure 2.1 Biological neuron

Figure 2.1 shows the structure of a neuron. The soma or cell body contains the metabolic machinery responsible for maintaining the integrity of the neuron. The nucleus contained in the soma is connected to a main connector known as the axon. The axon acts as an output channel for transmitting information from the neuron to other neurons. The dendrites are the branches leading to the neuron and act as input channels, gathering information from other neurons.

2.4 The Model of Artificial Neural Networks

2.4.1 McCulloch-Pitt Neuron

McCulloch-Pitts Neuron was developed by Warren S. McCulloch, a neuroscientist, and Walter Pitts, a logician in 1943. According to Faussett (1994), McCulloch-Pitts neuron is the earliest artificial neuron. McCulloch and Pitts tried to understand how the brain could produce highly complex patterns by using many basic cells that were connected together.

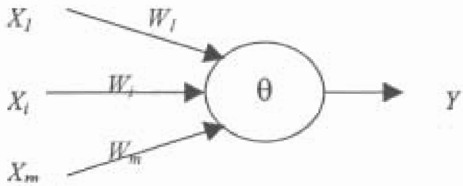


Figure 2.2 The McCulloch-Pitts neuron

2.4.2 The Perceptron

Frank Rosenblatt developed the Perceptron and the Perceptron theorem in 1950s. The Perceptron is a single-layer networks that is very similar to the McCulloch-Pitts Neuron except the bias θ is replace by a W_0. Georgiou (1997) said that the purpose of the Perceptron is to be used as a two-class classifier in which the input pattern is classified by adjusting the weights. The correct classification would only happened if they are linearly separable. The following function is used by the Perceptron as the node activation function.

$$\text{net} = \sum_{i=1}^{n+1} W_i X_i$$

$$f(\text{net}) = \begin{cases} 1 & \text{if net} >= 0 \\ -1 & \text{if net} < 0 \end{cases}$$

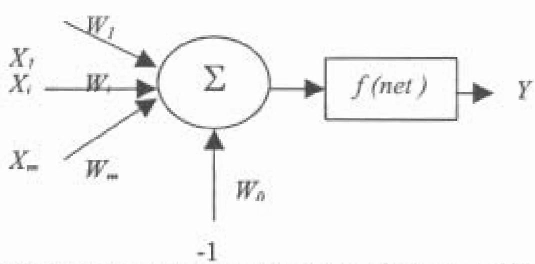


Figure 2.3 The Perceptron (modified from Georgiou, 1997)

2.5 Adaptive Resonance Theory

Adaptive Resonance Theory is developed by Carpenter and Grossberg in 1970s, using a mixture of interdisciplinary theory such as psychology, neurobiology and mathematics. There are various types of models in the ART family, both the supervised and unsupervised.

2.5.1 ART1

An unsupervised learning, self-organizing ART-based network architecture known as ART1 was introduced by Carpenter and Grossberg in 1987. ART1 had the capability of processing only binary input data.

2.5.1.1 ART Architecture

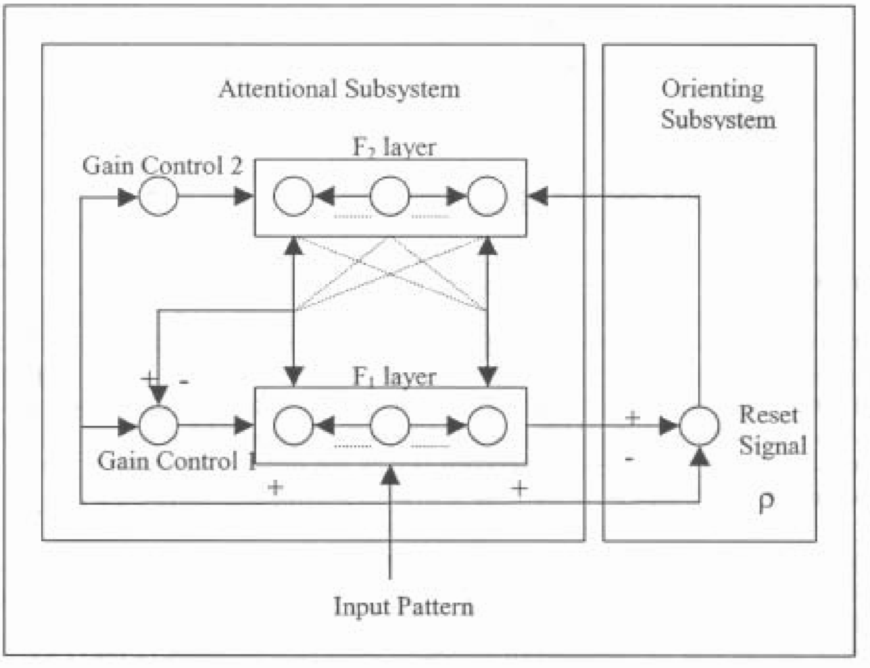


Figure 2.4 A generic architecture of an unsupervised ART network. (cited from Tan, 2001)

Figure 2.4 shows the generic architecture of ART. There are two layers of nodes in the network playing different roles at different times. The F_1 layer is the input comparison layer and the F_2 layer is the output or recognition layer. These layers are inter-linked by bi-directional feed forward and feedback connections. There are two sets of weights embedded in the connections that is the feed forward or bottom-up weights from F_1 to F_2 and the feedback or top-down weights from F_2 to F_1. There is some control logic handling the propagation of signals in the network, for instance the gain control and the reset circuit. According to Carpenter and Grossberg (1987a), F_1 and F_2 layer should meet a criterion called 2/3 (Two-Out-of-Three) rule for a node activation in F1 and F2 later.

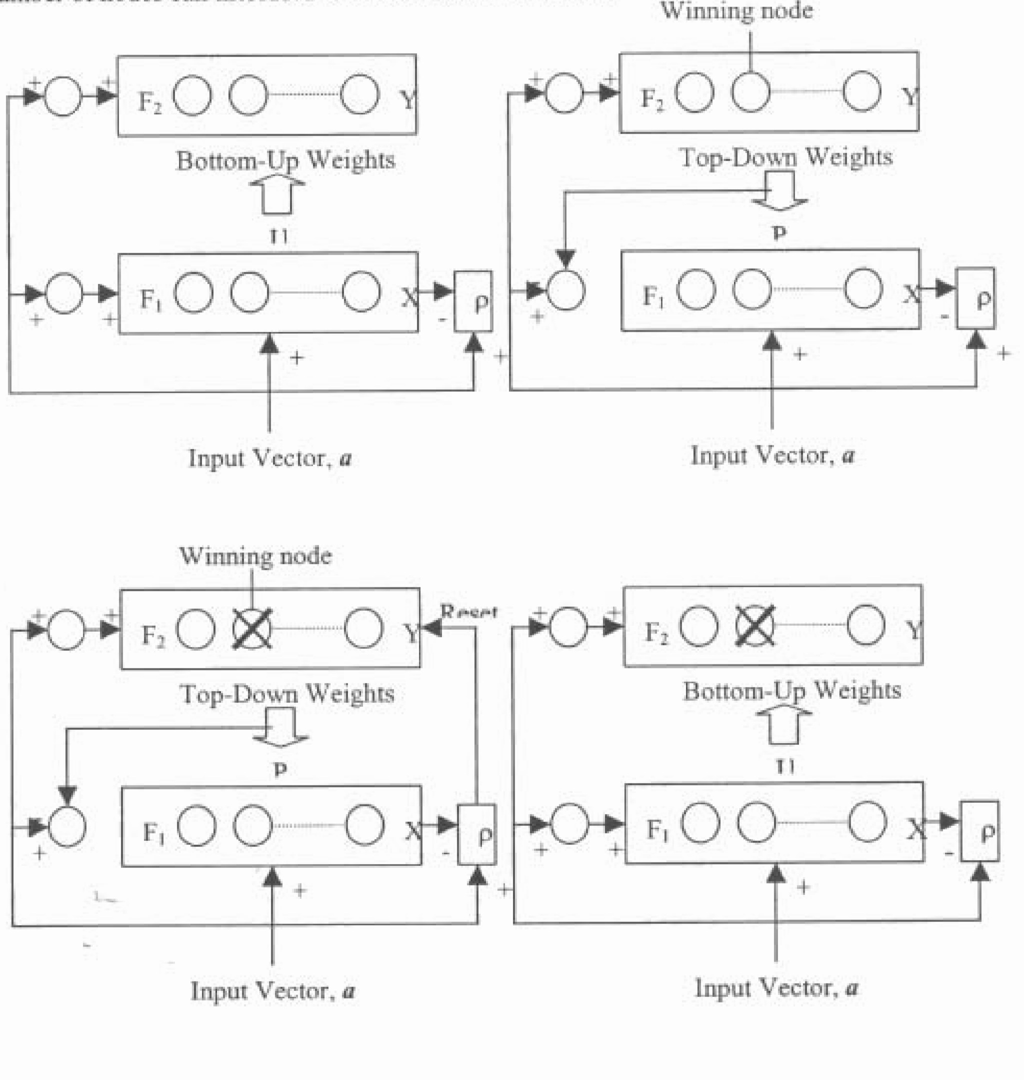
2.5.1.2 Pattern Matching Cycle

Figure 2.5 illustrates a typical matching cycle in ART. There are three cycles in the pattern matching cycles, which are the Hypothesis Selection, Hypothesis Test and Hypothesis Search.

In the Hypothesis Selection cycle, an input vector ***a*** is first registered itself as a pattern of Short-Term Memory (STM) activity X across the F_1 layer. The output pattern of X, U is transmitted from F_1 to F_2 through the adaptive bottom-up weights or the Long-term Memory traces. Each F_2 node receives the pattern U weighted by its corresponding stimulus. By the internal competitive dynamics of self-reinforcement and lateral inhibition or on-center off-surround competition (Carpenter & Grossberg, 1987a) in accordance with the winner-takes-all criterion, the node with the largest activation is selected as the winner whereas all other nodes are shut down. Thus only one component of Y corresponding to the winning node is non-zero.

During the Hypothesis Test cycle, the winning node reads out its prototype vector ***P*** to F_1 from the top-down weights. Eventually a new pattern of STM, X^* is induced across F_1. The pattern X^* and ***a*** are compared for their matching level at F_1. A vigilance test is carried out where the matching level of X^* with ***a*** is tested against a dimensionless threshold called the vigilance parameter, ρ, at the reset circuit. If the vigilance test is satisfied, the network enters a resonant state to allow the LTM traces of the winning node to learn or adapt to new information represented by the STM activity at F_1.

On the other hand in the Hypothesis Search, if the vigilance test fails, a search cycle is initiated to find F2 node that provides a better match with the input pattern. A reset signal is sent by the reset circuit to the F2 layer in order to refresh any activity in the network for a new hypothesis selection and test cycle. This reset signal also leads to the inhibition of the winning F2 node from participating in another pattern matching cycle. The search continues until an F2 node is able to satisfy the vigilance test. If no such node exists, a new node is created in F2 to code the input pattern. Therefore, F2 is a dynamic layer during learning process when the number of nodes can increased to absorb new information.



2.5.2 Fuzzy ART

Carpenter and Grossberg extended the ART1 network into Fuzzy ART network for fast-unsupervised learning and categorization of binary and analogue patterns. Fuzzy ART combines both the fuzzy set theory and Adaptive Resonance Theory for the exploitation of a close similarity between the computations of Fuzzy subsethood and the learning rules of ART1. In Fuzzy ART, the fuzzy AND and MIN operator replaced the crisp intersection operation that governs the dynamics of ART1. The MIN operator reduces to the intersection operator in binary cases, thus enabling Fuzzy ART to accept both binary and analogue patterns. Fuzzy ART and ART1, as shown in figure 2.4 have similar architecture and the same basic features of the hypothesis operation. However, Fuzzy ART only has only set of bi-direction weight vectors in which subsumes both the bottom-up and top-down weight vectors of ART1.

2.5.2.1 Fuzzy ART Algorithm

The dynamics of Fuzzy ART can be described in terms of the following four phases (Tan, 2002):

a) Initialisation

Each input pattern to the F_1 layer is an M-dimensional vector, $a=(a_1,\ldots,a_M)$, with each component $\boldsymbol{a}$ in the interval [0,1]. Each category node in the F_2 layer is linked by a vector of adaptive weights or the LTM traces, denoted as $W_j=(W_{j1},\ldots,W_{jM})$. At time t = 0, the weight vectors are initialised to unity,

$$W_{j1}(0) = \ldots = W_{jM}(0) = 1 \qquad j = 1,\ldots,N \qquad (2.1)$$

Each category is regarded as an uncommitted node. When learning takes places, an F_2 node becomes committed by modifying its weight vector to encode the input pattern. The resulting weight vector can be viewed as a category prototype of input patterns.

b) Category Choice, Test and Search

The input vector $\boldsymbol{a}$ is propagated to the F_2 layer. The response of each F_2 node is measured by using a choice function

$$T_j(a) = \frac{|a \wedge W_j|}{\alpha + |W_j|} \qquad (2.2)$$

By setting the $\alpha \approx 0$, the choice function defined by equation (2.1) reduces to a measure of degree of $\boldsymbol{a}$ being a fuzzy subset of W_j. For notation simplicity, $T_j(a)$ is written as T_j, since a is fixed. The node that has the highest response, denoted as node J is selected as the winning node,

$$T_J = max\{\, T_j: j=1,\ldots,N\} \qquad (2.3)$$

If there is a tie on T_j, then the category with the smallest index is selected. All other nodes $j \neq J$ are deactivated in accordance with the winner-takes-all competition.

The winning node J then propagates its weight vector back to F_1. A vigilance test is performed to calculated the similarity against the vigilance threshold between the transformed category prototype and the input vector:

$$\frac{|a \wedge W_J|}{|a|} \geq \rho \qquad (2.4)$$

If the vigilance test is satisfied, resonance occurs and learning takes place. Else is the test fails, node J is inhibited and it is prohibited from participating in subsequent competitions. Input a is re-propagated to F_2 for another cycle of searching for a new winner. The process is repeated, consecutively disable nodes in F_2, until either a category prototype is able to pass the vigilance test or if no such node exist, a new node is created to code the input vector.

c) Learning

Once search ends, learning takes place by adjusting the weight vector of the winning node according to

$$W_J^{(new)} = \beta(a \wedge W_J^{(old)}) + (1-\beta)W_J^{(old)} \qquad (2.5)$$

There are two learning modes: (1) fast learning corresponds to setting β=1 in which allows the weight vector to converge to the asymptotic category boundary in one attempt, (2) fast-commit and slow-recode learning corresponds to setting $\beta=1$ for an uncommitted node and $\beta<1$ for a committed node, in which slowly varies the weight vector to make the system more resistant to noise.

d) Complement Coding

Carpenter et al (1991a, 1992b) shows that an input pre-processing procedure that involves normalization of input patterns to retain the amplitude information,

$$|a| \equiv \gamma > 0 \qquad (2.6)$$

A normalization technique called complement coding is recommended. With complement coding, an M-dimensional input vector $\boldsymbol{a}$ is normalized to a $2M$-dimensional vector $\boldsymbol{A}$ as follows

$$A=(a,a^c)\equiv(a_1,\ldots,a_M, 1-a_1,\ldots,1-a_M) \qquad (2.7)$$

The magnitude of $\boldsymbol{A}$ is equal to $\boldsymbol{M}$ and it remains constant. When complement-coding is used, the number of nodes in F_1 is doubled, and the weight vector W_j defined by equation (2.1) extended to

$$W_{j1}(0)=\ldots=W_{j,2M}(0)=1 \qquad j=1,\ldots,N \qquad (2.8)$$

Thus the categories formed by Fuzzy ART are hyper-rectangles, in which each hyper-rectangle represents an interval of expected values for each input feature.

2.5.3 ARTMAP

ARTMAP, a type of supervised learning model from the ART family was introduced by Carpenter et al in 1991. Such a model is able to autonomously learn to classify arbitrarily many and arbitrarily ordered vectors into recognition categories based on predictive success. The ARTMAP system consists of a pair of ART modules, ART_a and ART_b, coupled via a map field. During the learning, ART_a module receives a stream of input patterns $\{a^{(p)}\}$ and ART_b receives a stream of input patterns $\{b^{(p)}\}$, where $b^{(p)}$ is the target classes of input. Each ART module self-organizes individually into categories representing the data or input pattern at ART_a and the supervisory signal or target or input pattern at ART_b. These modules are linked by an associative learning network and an internal controller that ensures autonomous system operation in real time. The map field however, does not directly associate $\boldsymbol{a}$ and $\boldsymbol{b}$ but by linking the compressed representation of the category prototypes at F_2^a layer to the target classes at F_2^b layer.

Similar to the ART modules, a control mechanism is also included in the map field to detect predictive error and to modulate learning. A vigilance test was carried out in the map field besides the individual vigilance test in ART_a and ART_b. This hypothesis testing determines whether or not the predicted class match as the actual class. An internal controller used conjointly maximizes predictive generalization and minimizes predictive error by linking predictive success to category size on a trail-by-trial basis using only local operation (Carpenter et al, 1991a). The ART_a vigilance parameter ρ_a is increased by a minimal amount needed to correct a predictive error at ART_b.

2.5.4 Fuzzy ARTMAP

Fuzzy ARTMAP is extended from ARTMAP by replacing the pair of ART modules with a pair of Fuzzy ART modules (Carpenter et al, 1991b). Hence the Fuzzy ARTMAP is a self-organizing, supervised learning model for classification of both analogue and binary patterns.

2.5.4.1 Fuzzy ARTMAP Algorithm

Figure 2.6 shows the architecture of the Fuzzy ARTMAP model. For clarity, the variables in ART_a and ART_b are designated by subscripts or superscripts a and b respectively and those of the map field are designated by a superscript ab.

ART_a and ART_b

Inputs to each Fuzzy ART module are in the complement code form:

For ART_a: input vector, $\boldsymbol{A}=(\boldsymbol{a},\boldsymbol{a}^c)$

F_1^a output vector, $\boldsymbol{X}_a = (x^a{}_1, \wedge, X^a{}_{2Ma})$

Weight vector, $W_j^a = (W^a{}_{j1},\wedge,W^a{}_{j,2Ma})$ $\quad j=1,\ldots,N_a$

For ART_b: target vector, $\boldsymbol{B}=(\boldsymbol{b},\boldsymbol{b}^c)$

F_1^b output vector, $\boldsymbol{X}_b = (x^b{}_1,\wedge,X^b{}_{2Mb})$

Weight vector, $W_k^b = (W^b{}_{k1},\wedge,W^b{}_{k,2Mb})$ $\quad k=1,\ldots,Nb$

Map field: output vector, $X_{ab} = (X_1{}^{ab}, \wedge , X_{Nb}{}^{ab})$

Weight vector from the j-th $F_a{}^2$ node to F^{ab},

$W_j^{ab} = (W_{j1}{}^{ab}, \wedge , W^{ab}{}_{jNb})$ $\quad j=1,\ldots,N_a$

Vector $\boldsymbol{X}^a$, $\boldsymbol{X}^b$, $\boldsymbol{Y}^a$, $\boldsymbol{Y}^b$ and $\boldsymbol{X}^{ab}$ are reset to 0 between input presentations.

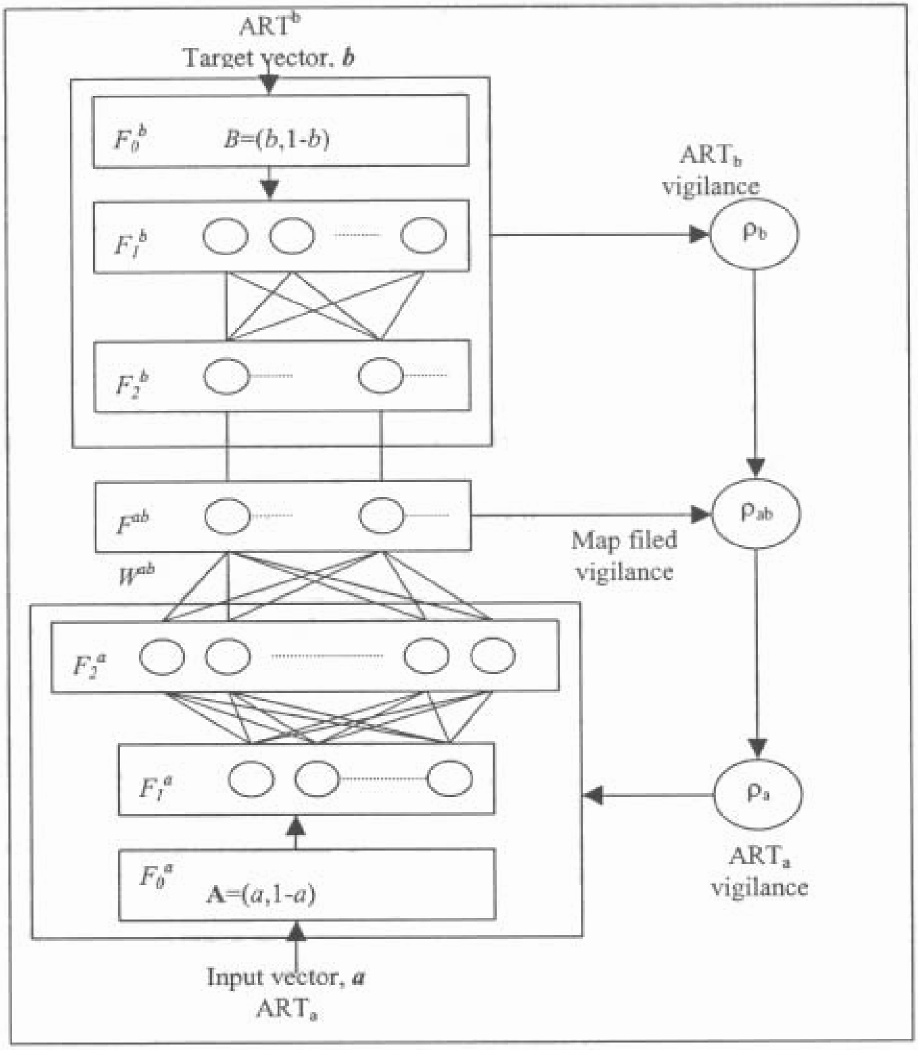


Figure 2.6 Fuzzy ARTMAP Architecture (cited from Tan, 2001)

There are four processes involve in the map field that is the map field initialisation, map field activation, match tracking and map field learning.

a) Map Field Initialisation

There is a one-to-one links between F^{ab} and F_2^b, thus the number of nodes in F_2^b and F^b is the same. The F_2^a and F_2^b is linked through an adaptive pathway routed by a set of weight vectors W^{ab}. Each component of the weight vector is initialised to 1 before the learning takes place. At time=0,

$$W^{ab}_{j1}(0)=K=W_{j,Nb}(0)=1 \qquad j=1,K,N_b \qquad (2.9)$$

Such setting indicates that the j-th prototype node in F_2^a is linked to every possible target class in F_2^b.

b) Map Field Activation

Once a pair of input-target vectors is present, the input vector is organized by ART_a by using the Fuzzy ART algorithm to search for the winning node J. Node J then fires F^{ab} by sending a prediction through W_J^{ab}. If K is the target class, then the output vector of F_2^b would be

$$Y_k^b = \begin{cases} 1 & \text{If } k=K \\ 0 & \text{otherwise} \end{cases} \qquad k=1,K,N_b \qquad (2.10)$$

Since the F_2^a and F_2^b are always active for supervised learning to occur in Fuzzy ARTMAP model, the F^{ab} output vector X^{ab} adheres to the following equation

$$X^{ab}=Y^b \wedge W_J^{ab} \qquad (2.11)$$

A map field vigilance test is carried out to confirm the prediction by measuring the similarity between the predicted vector X^{ab} and the target vector Y^b as follows

$$\frac{|X^{ab}|}{|Y^b|} = \frac{|Y^b \wedge W_J^{ab}|}{|Y^b|} \geq \rho_{ab} \qquad (2.12)$$

where $\rho_{ab} \in [0,1]$ is a user-defined map field vigilance parameter $0<\rho_{ab}<=1$. $|Y_b|$ always equals to 1. $|X^{ab}|$ can equals to either 1 if a correct prediction is made by node J, or 0 otherwise, by referring to equations (2.10) and (2.11). If node J has made a correct prediction, the learning in ART and the map field would take place. Otherwise a series of corrective actions known as match tracking is executed.

c) Match Tracking

When a predictive failure occurs at the map field, ρ_a is raised by a minimum value,

$$\rho_a = \frac{|A|}{|A \wedge W_J^a|} + \delta$$

where $0<\delta<<1$. This would result in a failure of the ART_a vigilance test because $|A \wedge W_J^a| < \rho_a|A|$. Hence a new winner in F_2^a must be found. In other words, match tracking provides a mean of selecting a prototype node, which satisfies both the ART_a and map field vigilance test.

d) Map Field Learning

An association is made upon the existence of the active J-th node of F_2^a and the active K-th mode of F_2^b. During the map field learning W_J^{ab} is updated to

$$W_{Jk}^{ab} = \begin{cases} 1 & \text{if } k=K \\ 0 & \text{if } k\neq K \end{cases}$$